

11-1500 A

IN THE UNITED STATES  
PATENT AND TRADEMARK OFFICE

11/14/11 JC951 U.S. PTO

PATENT APPLICATION

P K Chidambaran  
Thomas A Hoch  
John Patrick Jones  
Andrew A Long  
Prasasth R Palnati  
Ronald M Parker  
Raymond J Schmidt

"Express Mail" mailing label  
number EJ049977644 US  
Date of Deposit  
I hereby certify that this APPLICATION is being deposited  
with the United States Postal Service "Express Mail Post  
Office to Addressee" service under 57CFR 1.10 on the date  
indicated above and is addressed to the Commissioner of  
Patents and Trademarks  
Washington, D.C. 20231  
R. Stone  
(Printed name of person mailing paper of fee)  
R. Stone  
(Signature of person mailing paper of fee)

JC825 U.S. PTO  
09/11/1997

CASE 1-1-1-1-1-1-1

TITLE APPARATUS AND METHOD FOR REDUNDANCY OF  
PROCESSING MODULES INTERFACED TO A SWITCHING CORE

ASSISTANT COMMISSIONER FOR PATENTS  
WASHINGTON, D.C. 20231

SIR:

NEW APPLICATION UNDER 37 CFR § 1.53(b)

Enclosed are the following papers relating to the above-named application for patent:

Specification – 20 pages  
5 Informal Sheets of drawing(s)  
1 Assignment(s) with Cover Sheet – 9 pages  
Declaration and Power of Attorney – 9 pages

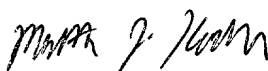
CLAIMS AS FILED				
	NO. FILED	NO. EXTRA	RATE	CALCULATIONS
Total Claims	25 - 20 =	5	x \$18 =	\$90
Independent Claims	3 - 3 =	0	x \$80 =	\$0
Multiple Dependent Claims, if applicable			+ \$270 =	\$0
Basic Fee				\$710
			<b>TOTAL FEE</b>	<b>\$800</b>

Please file the application and charge **Lucent Technologies Deposit Account No. 12-2325** the amount of \$800, to cover the filing fee. Duplicate copies of this letter are enclosed. In the event of non-payment or improper payment of a required fee, the Commissioner is authorized to charge or to credit **Deposit Account No. 12-2325** as required to correct the error.

The Assistant Commissioner for Patents is hereby authorized to treat any concurrent or future reply, requiring a petition for extension of time under 37 CFR § 1.136 for its timely submission, as incorporating a petition for extension of time for the appropriate length of time if not submitted with the reply.

Please address all correspondence to **Docket Administrator (Room 3C-512), Lucent Technologies Inc., 600 Mountain Avenue, P.O. Box 636, Murray Hill, New Jersey 07974-0636**. However, telephone calls should be made to me at 732-949-9742.

Respectfully,



Matthew J Hodulik  
Reg. No. 36164  
Attorney for Applicant(s)

Date: November 14, 2000

**Lucent Technologies Inc.**  
**600 Mountain Avenue (Room 3C-512)**  
**P.O. Box 636**  
**Murray Hill, New Jersey 07974-0636**

## APPARATUS AND METHOD FOR REDUNDANCY OF PROCESSING MODULES INTERFACED TO A SWITCHING CORE

### FIELD OF THE INVENTION

The present invention relates generally to communication systems and more particularly to packet switching systems having redundancy protection.

### BACKGROUND OF THE INVENTION

Multiservice switches used, for example, by communications providers in wide area networks typically provide a number of different interfaces for incoming and outgoing communications traffic to the core switching fabric in order to accommodate customer needs. These interfaces can range, for example, from high rate optical trunking ports to lower rate electrical interfaces. In general, the different interfaces are provided through service specific equipment grouped together on what are termed “service shelves”, where the service shelves then couple to the switching core. A typical service shelf will include the physical layer interface which couples to higher layer service cards (e.g. layer 2 or 3 for ATM or IP) and then to the switching core. Failure protection of equipment utilized in multiservice switches usually in the form of redundant circuit paths is also extremely important in order to provide the type of reliability that is necessary for these switches. Accordingly, extra service cards (or protection cards) are often provided within a service shelf to allow for the required protection.

In prior art multiservice switches of the type described above, the service cards from the service shelves are directly connected either logically or physically to a core port on the switching core. In accordance with this type of connection, a service card failure makes the core port(s) having the failed card thereon unavailable to the user which in turn degrades overall system bandwidth. The manner in which prior art systems provide service card redundancy is considered a major limitation in that a significant amount of core bandwidth is wasted upon failure of a service card. In addition, core bandwidth is wasted by the need to dedicate additional core ports in order to provide redundancy. Accordingly, there is a need to preserve maximum overall system bandwidth

in a packet switching system while at the same time incorporating protection of various service modules.

### **SUMMARY OF THE INVENTION**

The present invention is a core interface mechanism that permits 1:N type port protection on the core side of the switch such that core bandwidth is not wasted by the direct connection of service cards to the switching core. In an exemplary embodiment, a core interface module supports up to two active service cards and one dedicated protection service card. To provide increased efficiency and lower cost the redundant service card does not strand user bandwidth in the switch core. In an exemplary embodiment, the core interface includes a plurality of core side input and output ports for coupling to the switching core and a plurality of card side input and output ports for coupling to the service cards. A data flow switch function couples between the core side ports and the card side ports. The data flow switch function operates to complete data flow paths between the core side ports and the card side ports. A link failure detector detects a communications failure on a communications link associated with one of the active service cards and alters the data flow switch function such that one or more of said data flow paths associated with one of the active service cards are switched to one of said protection cards. The bandwidth allotment between the service cards and the interface device is generally greater than available bandwidth between the switching core and the interface device, thereby enabling connection of the protection cards without corresponding usage of switching core bandwidth.

### **BRIEF DESCRIPTION OF THE DRAWINGS**

A more complete understanding of the present invention may be obtained from consideration of the following detailed description of the invention in conjunction with the drawing, with like elements referenced with like references, in which:

FIG. 1 is a high level diagram of a multiservice switch incorporating the core interface device of the present invention;

FIG. 2 is an exemplary block diagram of a high speed service shelf in accordance with the present invention;

FIG. 3 is an exemplary block diagram of a core interface card for a high speed shelf;

FIG. 4 is an exemplary block diagram of an aggregator function as used in connection with the present invention;

5 FIG. 5 is an exemplary embodiment of a core interface card for a low speed shelf;

FIG. 6 is an exemplary embodiment of a higher level service card as used in connection with the present invention; and

FIG. 7 is an exemplary block diagram of an arbiter function as used in connection with the present invention.

## 10 **DETAILED DESCRIPTION**

Multiservice switches used by communications providers for wide area networks typically provide a number of different interfaces for access to and from the core switching fabric in order to accommodate customer needs. As discussed in the background, the different interfaces may be provided through service shelves which then  
15 couple to the switching core.

The present invention is a core interface mechanism that permits 1:N type port protection on the core side of the switch such that core bandwidth is not wasted by the direct connection of service cards to the switching core. Referring to Fig. 1, there is shown one exemplary embodiment of a multiservice switch 10. The switch includes a  
20 service shelf 12 which incorporates the core interface 14 of the present invention. As would be understood, the functional blocks illustrated in the figure may take the form of one or more cards or modules that are insertable into a rack or other similar type system. The service shelf 12 couples to first and second redundant switching cores 16, 18. A second service shelf 20 couples to what can be considered the output side of the  
25 switching cores.

As shown, the general makeup of the service shelf 12 includes a physical layer interface card 22 which is a user interface that can be an optical or electrical interface, e.g., DS3, OC-12, OC-48, OC-192, etc. In the case of the high speed shelf shown, the physical layer is generally a high density optical interface such as OC-48 or OC-192.

30 The physical layer card 22 couples to higher level service cards 24, 26 (for example,

layer 2 or layer 3 for ATM or IP) through a cross connect device, for example, a SONET STS-1 level cross-connect. The service cards 24 couple to the switching core through core interface modules 14. As shown, the switching cores 16, 18 are traditional switch cores including input/output ports 32 as well as switching fabrics 34.

5 The interface mechanism between the service cards 12 and the core 16, 18 provides redundancy protection between the service cards and core without the requirement that extra core bandwidth be allotted for the protection cards. As shown in the exemplary embodiment, two on-line ATM service cards 24 are protected by one back-up or protect service card 26. The core interface card 14 permits routing of core data to and from any of the three cards. In addition, the protection card 26 can be  
10 switched in place without the corresponding re-routing having to be known to the rest of the system.

Referring to Fig. 2, a detailed block diagram of a service shelf 12 in accordance with the present invention is shown. Fig. 2 illustrates the interface between the service  
15 cards 24, 26 and the switching core via the core interface modules 14, where the specific interconnects between the service cards and the core interface are shown. In the exemplary embodiment, the service shelf 12 includes nine service cards (SC0 – SC8) which couple, respectively, to six core interface cards (CI0 – CI5). As in Fig. 1, two on-line service cards 24 and one protect service card 26 couple to the switching cores  
20 through each core interface card providing 1:2 redundancy. Also included in the service shelf are shelf control processor cards 36 which handle administrative processing functions for the shelf.

The core interface cards 14 couple to redundant switch cores 16, 18. A core interface card 14 monitors its link to the core and reports status to the shelf control  
25 processor 36 on the service shelf. Referring to Fig. 3 in combination with Fig. 2, an exemplary block diagram of a core interface card 14 is shown. As shown, service cards 24, 26 couple to the core through an aggregator device 38 in the core interface card 14. Interconnections between the aggregator in the core interface and the arbiter blocks on the service cards are illustrated with double arrows. (Fig. 2).

30 The aggregator device 38 acts as an interface between the service cards 12 and the switching core and essentially distributes core traffic throughout the service shelf. The

aggregator 38 acts a datapath flow switch, directing flows to either the normally active service card slot or to the dedicated protection slot. In all cases, the aggregator 38 will allow control information connectivity through the core to all attached service cards 24, 26 and shelf control processors 36. Although shown and described as an applications specific integrated circuit (ASIC), it would be understood that the functionality of the aggregator 38 as described herein may also be implemented using discrete components. As shown in Fig. 2 and 3, the core side of the aggregator 38 couples to multiple serializer/deserializer blocks 40. The implementation and function of a serializer/deserializer would be well known to a person skilled in the art. The serializer/deserializers 40 couple to optical/electrical (O/E) components 42 in order to provide the interface to the switching core. Failure of a link will be detected by a serializer/deserializer 40 or the aggregator device 38 and reported to the shelf control processor 36 through a control interface on the aggregator. Failures may be detected, for example, by the loss of a clock signal corresponding to the link or an invalid parity across the link. Other types of failures that are detectable and that can be characterized as a link failure would be apparent to those skilled in the art. As will be explained, the shelf control processor 36 (in combination with the aggregator 38) trigger appropriate corrective action in response to a link failure. The aggregator 38 on the core interface card 14 also contains a thread switch function 44 for service card protection. The switch function 44 allows the core interface card 14 to steer traffic on a given thread to/from an active service card to a protection card. For the shelf, service card protection will be 1:2. The core interface card 14 (and the shelf control processor 36) will control the protection switching of the interface. In addition, as will be explained, an arbiter function on the service card can detect link failures on the basis, for example, of the receipt/non-receipt of link test cells.

Fig. 4 shows a functional block diagram of the aggregator device 38. The aggregator 38 includes ingress receive logic 50 and egress transmit logic 52 on the service card side. Ingress transmit logic 54 and egress receive logic 56 are also found on the core side of the aggregator 38. There are two aggregation functions – AGR0 and AGR1 – implemented in the aggregator (AGR) ASIC, each performing an aggregation of up to 6 independent data streams into a 2.5Gbps thread. These two aggregation functions

are independent and the operation of one does not affect any state of the other. In one exemplary embodiment, each aggregator function AGR0, AGR1 includes a multiplexer unit 58 which couples to the ingress receive logic 50, a cell decode unit 60 which couples to the output of the multiplexer 58 and a buffer management unit 62 which couples to the output of the cell decode unit 60. A credit/grant manager function 64 and a multicast unit 66 each couple to the output of the buffer management unit 62. A virtual output queue (VOQ) memory interface 68 and a pointer memory interface 70 each couple to the multicast unit 66. A VOQ scheduler 72 couples to the credit/grant manager 64.

The AGR ASIC communicates with the service shelf cards through an arbiter (ARB) ASIC 76 over an 8-bit LVDS (low voltage differential signal) interface (Fig. 2). This interface runs, for example, at 266 MHz with data being transferred on both clock edges. As shown, the AGR ASIC has 8 ARB interface (AIF) ports. Four of these AIF ports can be configured to connect to either of the aggregation functions in the AGR ASIC. Of the remaining four AIF ports (P0 – P7), two are connected to aggregation function 0 (AGR0) and the other two are connected to aggregation function 1 (AGR1). Thus, a maximum of six AIF ports can be connected to each aggregation function. In the ingress direction, each aggregation function statistically multiplexes a combination (maximum of 6 data streams) of OC-12, 2 x OC-12, and OC-48c data streams into a 2.5 Gbps stream. In the egress direction, each aggregation function broadcasts an OC-48 thread coming from the core to the six (6) ARB ASICS connected to that thread.

As discussed above, the AGR ASIC communicates with the switch core, for example, on OC-48 links through quad serializer/deserializer (Serdes) 40 and Optical/Electrical ports 42. The Serdes transmitter 40 serializes and encodes the data, e.g. 8B10B data, for proper transmission over the fiber link. The receiver will deserialize, decode and also synchronize the four channels (channel lock) before transmitting the data to the aggregator (AGR) ASIC 38. Optical/Electrical components take the electrical signals produced by the Serdes and convert them to optical signals for fiber link transmission and take optical signals from the link and convert them to electrical signals for Serdes processing. In one embodiment of the invention, for example, a 96-byte data cell is striped among four channels. This data cell includes the 84-byte packet and 12-bytes of control data. Data is transferred between the aggregator



ASIC and each Serdes on a 4 x 8-bit unidirectional bus. This cell is transmitted in twenty-four 155.52 MHz-clock cycles.

The AGR ASIC 38 is used in high speed and low speed applications, where the respective service shelves are accordingly termed high speed service shelves (HSS) and low speed service shelves (LSS). In the HSS and LSS applications, the AGR 38 resides in the HSS and LSS core interface cards, respectively. In the exemplary embodiment of the high speed shelf 12, the core interface card in the HSS uses two AGR ASICS 38 and provides 10Gbps (4x2.5Gbps) interface to the switch core. In the exemplary embodiment of the low speed shelf (see Fig. 5), the core interface card 80 in the LSS uses one AGR ASIC 38 and provides 5Gbps (2x2.5Gbps) interface to the Switch Core. The AGR is software configurable based on the specific application.

In the exemplary embodiment, the AGR ASIC includes 8 AGR-ARB interfaces each with a data rate of OC-48. All of the eight AGR-ARB interfaces (AIF ports P0 through P7) are software configurable to operate the AGR ASIC in different configurations required for different shelves (e.g. the High-Speed Shelf and Low-Speed Shelf). Setting a corresponding port enable bit in *AIF Port Control Register 0 & 1* can activate each interface. AIF ports P0 & P1 are connected to the aggregation function 0 (AGR0) and ports P6 & P7 are connected to aggregation function 1 (AGR1). Ports P2 through P5 can be connected to either aggregation functions (AGR0 or AGR1), depending upon the *AGRn\_SEL* bit in the *AIF Port Configuration Register*. Therefore, at any time at most 6 AIF ports can connect to one OC-48 thread.

In the high-speed shelf, the core interface card 14 has two AGR ASICs 38 (AGR-A and AGR-B) residing on it and provides an aggregate bandwidth of 10Gbps to the core. Each AGR ASIC 38 is connected to one 5Gbps high-speed service card and to one of the two 2.5G ARB interfaces on the protection card. One of the two AGR ASICs will also have a shelf control processor (SCP) card(s) connected to it. Each SCP has an average data rate of 622Mbps (OC-12).

In the low-speed shelf (Fig. 5), the core interface card 80 has one AGR ASIC 38 and provides two 2.5Gbps aggregated threads to the core. The AGR ASIC interfaces with the ARB ASIC in 4 low-speed service cards, 2 protection cards, and 2 shelf control processor (SCP) cards. All low-speed cards have an average data rate of 2xOC-12,

however, in burst traffic conditions, the interfaces can support a peak data rate of OC-48. Fig. 5 shows AGR in LSS core interface card.

Referring again to Fig. 1 and 2, it can be seen that the service cards 24, 26 will receive flows from the redundant cores through the core interface card 14. An arbiter function (ARB) 76 in the service cards 24, 26 will monitor the end to end path of the flows through special in-band test messages over both cores. If a flow is failed, the destination ARB will automatically switch and accept traffic through the protection path from the redundant core and core interface card. The source ARB will always broadcast traffic and test messages through both cores. The AGR interfaces with an Arbiter device/circuit that resides on all service cards and shelf control processors 36 to complete the core interface. From a high level the ARB 76 is intended to merge traffic flows from each core as necessary, on a per flow basis, and act as a header translator and filter for traffic flows from the cores. The ARB and AGR will also provide flow checking and fault detection checking. A significant advantage of the present invention is the ability to switch individual flows without impacting other flows within the switching system.

Referring to Fig. 6, one exemplary embodiment of a high level service card 12 is shown. As illustrated, the service card is an ATM service card, although it would be understood that other types of service cards, for example IP, frame relay, and TDM. The service card shown provides 2 x 2.5 Gbps threads and provides the ATM layer and traffic management functions for the service shelf. As shown, cross connect interface terminations 86 couple to the ATM (layer 2) processing blocks 88. The ATM blocks 88 couple to respective traffic management functional blocks 90 as well as to the ARB ASIC 76 providing the two threads. The ATM layer blocks 88 also couple to a segmentation and reassembly function (SAR) 92 that couples to a local processor 94 via a PCI bus. The service card also includes timing and power functions 98.

The Arbiter ASIC, or ARB ASIC 76, will be used in the switching system as a flow control mechanism for cell traffic as well as a test cell generator and receiver for system level flow verification. As with the aggregator device, although the exemplary embodiment is described with respect to an ASIC, it would be understood that such a device may also be implemented using discrete components. The ARB is utilized, for example, in the high speed shelf, the low speed shelf, and interfaces on one side to a

physical layer device such as a scheduler, also known as a traffic manager or TM. On the opposite side, the ARB interfaces to the aggregator (AGR). The ARB ASIC includes a UTOPIA II bus for interfacing with a SAR for processor to processor communication. The ARB also supports an external memory interface for GMID (global multicast ID) to ECID (egress circuit ID) translation. The ARB ASIC contains a test cell generator and a test cell receiver to test online and off-line cell flows through the core.

The ARB resides on a service card and forwards user traffic (from the physical interface) to the core interface cards at an OC48 (2.5Gbps) rate. The ARB receives traffic from the core interfaces and will forward traffic destined to its TM device. An ARB also resides on the SCP. In the SCP application, the ARB interfaces to a SAR device to enable processor to processor communication and will not interface to a TM device.

Referring to Figure 7, a functional block diagram of the ARB ASIC 76 is shown. The exemplary embodiment of the ARB includes six interfaces: a PCI (processor interface) interface, a physical layer interface (PI Sched RX and TX), a SAR interface (RX and TX), two AGR interfaces (RX and TX, one per core) and an external memory interface. As discussed previously, the ARB includes a link test cell generator 102 and a link test cell receiver 104 which will be used in the system to verify flow integrity. The link test cell (LTC) generator 102 and receiver 104 couple to the aggregator interface 106, the link test cell receiver 104 coupling through respective egress filters 108. The ARB also includes internal priority queues (four QOS levels) 110 for egress traffic, the inputs of which couple to the egress filter 108. The priority queues couple to egress transmit ports (TM and Utopia) 112, 114 through a scheduler 116 or 118. The egress filters 108 in the ARB provide a filtering function that is used to determine if the ARB should accept unicast and multicast cells from the AGRs.

The ARB 76 operates in one of two modes. If the ARB resides on a service card (either in the high speed shelf or the low speed shelf), the ARB will be in TM mode in which all traffic is sent and received via the TM device or via the test cell interface. If the ARB resides on a processor card the ARB will be in SAR mode in which all traffic will be sent and received via the SAR or via the test cell interface.

From an ingress standpoint (with relation to the core), if the ARB 76 is in TM mode, user cells will enter through the physical layer interface TM. BIP8 calculations (bit interleaved parity across 8 bit boundaries) will be checked on a per cell basis and optionally drop BIP8 erred cells. Cells entering the ARB through the physical layer interface will be broadcast to both AGR ports (and sent to both cores). Internally generated link test cells will be combined with the user traffic in the ARB ASIC and sent to both AGR ports. The link test cell generator 102 can optionally back pressure the TM device using a back pressure table 116 to create space for test cell insertion. If no user cells or test cells exist, idle cells will be inserted to sustain the flow.

If the ARB is in SAR mode, cells will be accepted from the SAR device and the TM interface will be ignored. Again, the SAR cells will be combined with the internally generated test cells and sent to both AGR ports.

From an egress standpoint, cells will enter the ARB via one of two AGR interfaces. When a cell first enters the ARB, a check will be done to determine if the cell is a test cell, a unicast cell, a multicast cell, or an idle cell. Filters and checks will be done to forward the cell to the appropriate interface (TM/SAR or LTC receiver). BIP8 calculations will be checked on a per cell basis and optionally drop BIP8 erred cells. Cells destined for the TM/SAR are placed in one of four priority queues 110 based on a QOS field in the cell. Cells from both AGR interfaces are placed into the same queues. Cells will be read from the priority queues based on either a fixed priority or a programmable priority depending on scheduler mode and sent to the TM or SAR based on mode.

The egress queue back pressure mechanism will exist to prevent the egress priority queues from overflowing. Back pressure information will be inserted into the ingress path back to the AGRs. The ARB will also track and forward back pressure information from the AGRs to the TM device.

The PCI interface 120 provides access to on chip register and tables as well as off chip memory. In an exemplary embodiment, the PCI interface will be 32 bits wide and support a maximum frequency of 33 MHz. Burst access will be provided to on chip tables and off chip memory when the corresponding function is not enabled.

In accordance with the present invention, it can be seen that at the service shelf (SS) level the core interface cards are redundant on a per core basis. The service cards (SC) are 1:N redundant, (e.g., 1:2) without wasting core bandwidth. The AGR provides support for 1:N service card redundancy in the HSS and LSS applications. Figure 3 and Figure 5 depicted the AGR in the high-speed and the low-speed configurations. In the HSS application (see Figure 3), the core interface card 14 connects to one protection card (PC) that can protect any one of two service cards (SC0 and SC1). In the LSS application (see Figure 5), the core interface card 80 connects to two protection cards (PC0 and PC1) each can protect any of the four service cards (SC0, SC1, SC2, and SC3).

In the HSS application, as shown in Fig. 3, ARB0 and ARB1 of the SC0 and ARB0 of the PC are connected to AGR ASIC-A. Similarly, ARB0 and ARB1 of the SC1 and ARB1 of the PC are connected to AGR ASIC-B. Since there are two service cards (SC0 and SC1) each connected to two different AGR ASICs and there is only one protection card (PC) to protect them, a cross-connect is needed between two AGR ASICs on the HSS CIC card. When PC is protecting the SC0, PC-ARB0 protects SC0-ARB0 directly and PC-ARB1 protects SC0-ARB1 indirectly through the external cross-connect. Conversely, when PC is protecting the SC1, PC-ARB1 protects SC1-ARB1 directly and PC-ARB0 protects SC1-ARB0 indirectly through the external cross-connect. The cross-connect enable bit (*XCON\_EN*) in the *AIF Redundancy Register* is provided to enable and disable the external cross-connect. When enabled, the protection port on the AGR ASIC protects the “remote” ARB connected through the external cross-connect. When *XCON\_EN* is disabled, the protection port on the AGR would protect the “local” ARB. For example, if the *XCON\_EN* bit in AGR ASIC A is enabled, PC-ARB0 would protect SC1-ARB0 through the external cross-connect. If the *XCON-EN* bit in AGR ASIC A is disabled, PC-ARB0 would protect SC0-ARB0. This *XCON-EN* bit is used in HSS applications only and it should be disabled in LSS and NEP applications.

In the LSS application (see Figure 5), since there is only one AGR ASIC on the core interface card, external cross-connect is not needed. Therefore, the *XCON\_EN* bit is disabled and only *AGRn\_SEL* bits for the protection ports are used to configure the protection ports. On the ingress side, data from a protection card can go to one of two OC-48 threads to the switch based on the card it is protecting. Similarly, on the egress

side, data from one of two threads can now go to a protection card. The *AGR<sub>n</sub>\_SEL* bit (in *AGR<sub>n</sub> Port Configuration Register*) associated with the protection port is used to select one of two threads. This bit is set by the processor during switchover.

5 As discussed, support for 1:N service card redundancy is provided in the AGR 38. In the described embodiments of the HSS and the LSS one protection card (a hot standby) is provided for every two service cards. In order to provide the redundancy protection and allow for seamless traffic switchover between the protection card and service card, an address mapping scheme, termed a Z-mapping scheme (after the different  
10 address fields) is implemented.

All the ARB ASICS 76 in a switch utilizing the present invention interface are uniquely identified from a flow/connection standpoint based on an X.Y.Z addressing scheme. The X portion of the address represents an 8-bit OC192 port ID used for addressing one of 256 fabric output ports. A 2-bit Y field addresses the four OC 48 ports  
15 within an OC 192 port addressed by X. That is, Y specifies one of the four OC48 links between the switching core and a core interface card. A 3-bit Z field addresses an ARB ASIC or AIF port associated with an OC48 thread (PIF thread). The X.Y.Z value is stored in the packet header and is used by the switch fabric in the core and the line card on the service shelf to route packets to the correct destination card/port.

20 On the egress side, all user data cells and test cells received from the core are broadcast to all ARBS associated with an OC48 PIF thread. These cells contain a 3-bit E\_Z (egress) field that identifies one of 8 destination ARBs connected to the AGR. Each ARB also has a unique Z ID stored in its Z[2:0] register. Upon receiving a cell from the AGR, the ARB compares the E\_Z[2:0] field of the incoming cell with its Z ID. If the Z  
25 values match, the cell is processed, otherwise the cell is dropped.

When a service card fails, the associated egress traffic is switched to a protection card. In order to accomplish the switching, the AGR uses a 3-bit wide, eight entry deep Z-mapping table with each entry associated with one of the eight AIF ports. Each entry in the Z-mapping table contains the current mapped/unmapped Z address of the  
30 corresponding AIF port. The egress transmit logic in the AGR receives a cell from the egress receive logic, it looks up the Z mapping table used to overwrite the E\_Z field of

the outgoing egress cell. During normal operation, each entry in this table contains the Z address of the ARB connected to the associated AIF port. When one of the service cards fails, the Z address of the failed card and the protection card are swapped by the associated software. The Z address of the failed service card is now mapped to the Z address of the protection card and vice versa. Consequently, the egress traffic destined for the failed service card will now be accepted by the protection card.

It is desirable to have the Z-mapping table lookup disabled for test cells. For example, when a service card is being protected, it must still be able to receive test cells destined to it. Thus, test cells destined for the failed service card must not be mapped whereas user data cells destined for the same card must be mapped. The IGNR\_Z bit in the egress cell header is therefore provided to override the Z-mapping lookup table. Hence, the Z-mapping table lookup will only be performed when the IGNR\_Z bit is set to 0.

The foregoing description merely illustrates the principles of the invention. It will thus be appreciated that those skilled in the art will be able to devise various arrangements, which, although not explicitly described or shown herein, embody the principles of the invention, and are included within its spirit and scope. Furthermore, all examples and conditional language recited are principally intended expressly to be only for instructive purposes to aid the reader in understanding the principles of the invention and the concepts contributed by the inventor to furthering the art, and are to be construed as being without limitation to such specifically recited examples and conditions. Moreover, all statements herein reciting principles, aspects, and embodiments of the invention, as well as specific examples thereof, are intended to encompass both structural and functional equivalents thereof. Additionally, it is intended that such equivalents include both currently known equivalents as well as equivalents developed in the future, i.e., any elements developed that perform the same function, regardless of structure.

In the claims hereof any element expressed as a means for performing a specified function is intended to encompass any way of performing that function including, for example, a) a combination of circuit elements which performs that function or b) software in any form, including, therefore, firmware, microcode or the like, combined with appropriate circuitry for executing that software to perform the function. The

invention as defined by such claims resides in the fact that the functionalities provided by the various recited means are combined and brought together in the manner which the claims call for. Applicant thus regards any means which can provide those functionalities as equivalent as those shown herein. Many other modifications and applications of the principles of the invention will be apparent to those skilled in the art and are contemplated by the teachings herein. Accordingly, the scope of the invention is limited only by the claims appended hereto.



**What is claimed is:**

1           1.       An interface device for coupling a group of service cards of a packet based  
2 switch to a switching core thereof, said group of service cards including one or more on-  
3 line service cards and one or more protection cards, said device comprising:

4               a plurality of core side input and output ports for coupling to said switching core  
5 and a plurality of card side input and output ports for coupling to said service cards;

6               a data flow switch function coupled between said core side ports and said card  
7 side ports, said data flow switch function operable to complete data flow paths between  
8 said core side ports and said card side ports; and

9               a link failure detector operable to detect a communications failure on a  
10 communications link associated with one of said on-line service cards, said link failure  
11 detector further operable to alter said data flow switch function such that one or more of  
12 said data flow paths associated with one of said on-line service cards are switched to one  
13 of said protection cards.

1           2.       The device of Claim 1, wherein a bandwidth allotment between said  
2 service cards and said interface device is greater than available bandwidth between said  
3 switching core and said interface device, thereby enabling connection of said protection  
4 cards without corresponding usage of switching core bandwidth.

1           3.       The device of Claim 1, further including an aggregation function coupled  
2 between said card side ports and said core side ports, said aggregation function operable  
3 to combine multiple data streams into larger bandwidth threads.

1           4.       The device of Claim 3, wherein said data flow switch function switches  
2 threads between said on-line cards and said protection card.

1           5.       The device of Claim 1, further including a digital processor coupled to  
2 said link failure detector, said digital processor being coupled to said data flow switch  
3 function to control the activation thereof in response to detection of a link failure.

1           6.       The device of Claim 5, wherein each of said service cards is assigned an  
2 address, each packet traversing said switch also having an address, wherein upon  
3 detection of a link failure, said processor is operable to swap the address of one or more  
4 of said ports assigned to a failed on-line service card with that of said ports assigned to a  
5 protection service card to thereby accomplish data flow switching to said protection card.

1           7.       The device of Claim 6, wherein said data flow switching is accomplished  
2 on a per flow basis, data flow switching of one particular flow not affecting other  
3 different flows in said switch.

1           8.       The device of Claim 6, wherein said address swap of on-line service card  
2 and protection card is accomplished in a look-up table.

1           9.       The device of Claim 6, wherein said address is a hierarchical address  
2 scheme having multiple fields pertaining to various type links within said interface.

1           10.      The device of Claim 1, wherein said link failure detector includes a  
2 monitor to detect failures selected from the group consisting of bit interleaved parity,  
3 parity, and CRC.

1           11.      The device of Claim 1, wherein said link failure detector includes a  
2 receiver for monitoring per flow test cells.

1           12.      The device of Claim 1, wherein said service cards are layer 2 or layer 3  
2 cards.

1           13.      The device of Claim 12, wherein said service cards are selected from the  
2 group consisting of ATM, IP frame relay and TDM.

1           14.      The device of Claim 12, wherein said service cards include traffic  
2 management functionality.

1           15.     The device of Claim 3, wherein said interface device includes at least one  
2     serializer/deserializer coupled to said aggregator.

1           16.     An interface device for reducing the amount of core bandwidth necessary  
2     to be allotted to support both on-line service cards and back up protection cards in a  
3     multiservice switch, said device being coupled between said service cards and said  
4     switching core, said device comprising:

5                 a first plurality of input and output ports for coupling to both said on-line service  
6     cards and said protection cards;

7                 a second plurality of input and output ports coupled to said switching core;

8                 a data flow switch function coupled between said first plurality of input and  
9     output ports and said second plurality of input and output ports;

10                a link failure detector also coupled between first plurality of input and output  
11    ports and said second plurality of input and output ports; and

12                a controller coupled to said link failure detector and said data flow switch  
13    function, said controller operable to switch a data flow route between said service cards  
14    and said switching core from one of said on-line service cards to one of said protection  
15    cards upon detection of a link failure by said link failure detector.

1           17.     The device of Claim 16, wherein a bandwidth allotment between said  
2     service cards and said interface device is greater than the available bandwidth between  
3     said switching core and said interface device, thereby enabling connection of said  
4     protection cards without corresponding usage of switching core bandwidth.

1           18.     The device of Claim 16, further including an aggregation function coupled  
2     between said first plurality of ports and said second plurality of ports, said aggregation  
3     function operable combine multiple data streams into larger bandwidth threads.

1           19.     The device of Claim 16, said data flow switch function switches threads  
2 between said on-line cards and said protection card.

1           20.     The device of Claim 16, wherein said data flow switching is accomplished  
2 on a per flow basis, data flow switching of one particular flow not affecting other  
3 different flows in said switch.

1           21.     The device of Claim 16, wherein each of said service cards is assigned an  
2 address, each packet traversing said switch also having an address, wherein upon  
3 detection of a link failure, said processor is operable to swap the address of one or more  
4 of said ports assigned to a failed on-line service card with that of said ports assigned to a  
5 protection service card to thereby accomplish data flow switching to said protection card.

1           22.     The device of Claim 21, wherein said address swap of on-line service card  
2 and protection is accomplished in a look-up table.

1           23.     The device of Claim 21, said address is a hierarchical address scheme  
2 having multiple fields pertaining to various type links within said interface.

1           24.     A method of coupling equipment specific service cards to a switching core  
2 in a multiservice switch in order to conserve switching core bandwidth, said method  
3 comprising the steps of:

4           coupling on-line and protection service cards to said switching core through a  
5 core interface device, said core interface device including a data flow switch function for  
6 coupling service card side ports with core side ports thereof, said core interface device  
7 providing a given redundancy ratio of protection cards to on-line service cards;

8           detecting a link failure associated with an on-line protection card; and

9           altering said data flow function in response to detection of said link failure to  
10 provide one or more data flow paths from said on-line service cards to said protection  
11 service cards.

- 1            25.    The device of Claim 24, wherein a bandwidth allotment between said  
2    service cards and said interface device is greater than the available bandwidth between  
3    said switching core and said interface device, thereby enabling connection of said  
4    protection cards without corresponding usage of switching core bandwidth.

## **ABSTRACT**

The present invention is a core interface mechanism that permits 1:N type port protection on the core side of the switch such that core bandwidth is not wasted by the direct connection of service cards to the switching core. In an exemplary embodiment, a core interface module supports up to two active service cards and one dedicated protection service card. To provide increased efficiency and lower cost the redundant service card does not strand user bandwidth in the switch core. In an exemplary embodiment, the core interface includes a plurality of core side input and output ports for coupling to the switching core and a plurality of card side input and output ports for coupling to the service cards. A data flow switch function couples between the core side ports and the card side ports. The data flow switch function operates to complete data flow paths between the core side ports and the card side ports. A link failure detector detects a communications failure on a communications link associated with one of the active service cards and alters the data flow switch function such that one or more of said data flow paths associated with one of the active service cards are switched to one of said protection cards. The bandwidth allotment between the service cards and the interface device is generally greater than available bandwidth between the switching core and the interface device, thereby enabling connection of the protection cards without corresponding usage of switching core bandwidth.

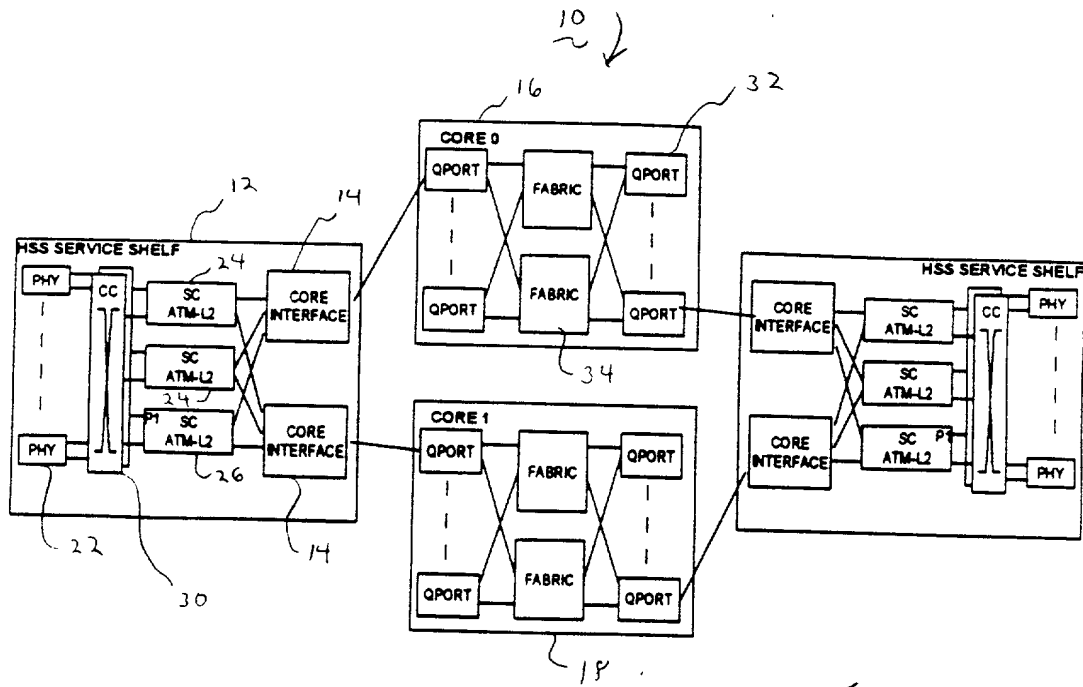


Fig 1

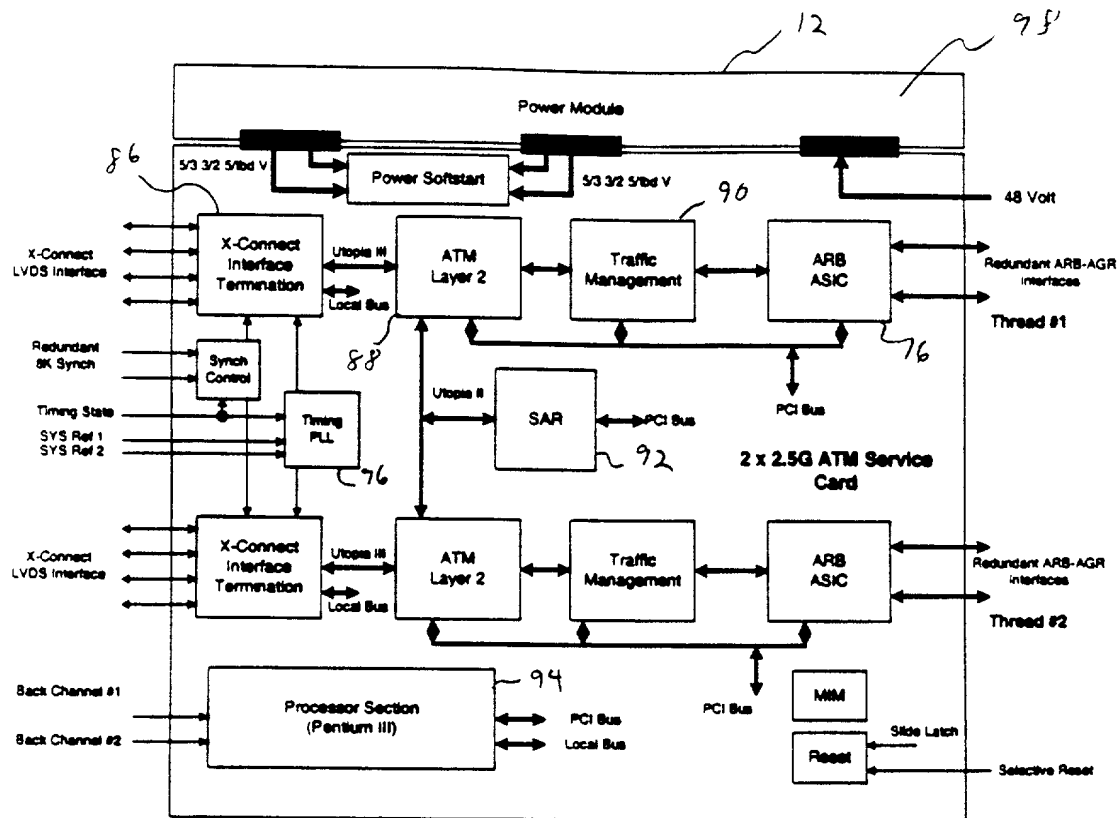


Fig 6

2/5

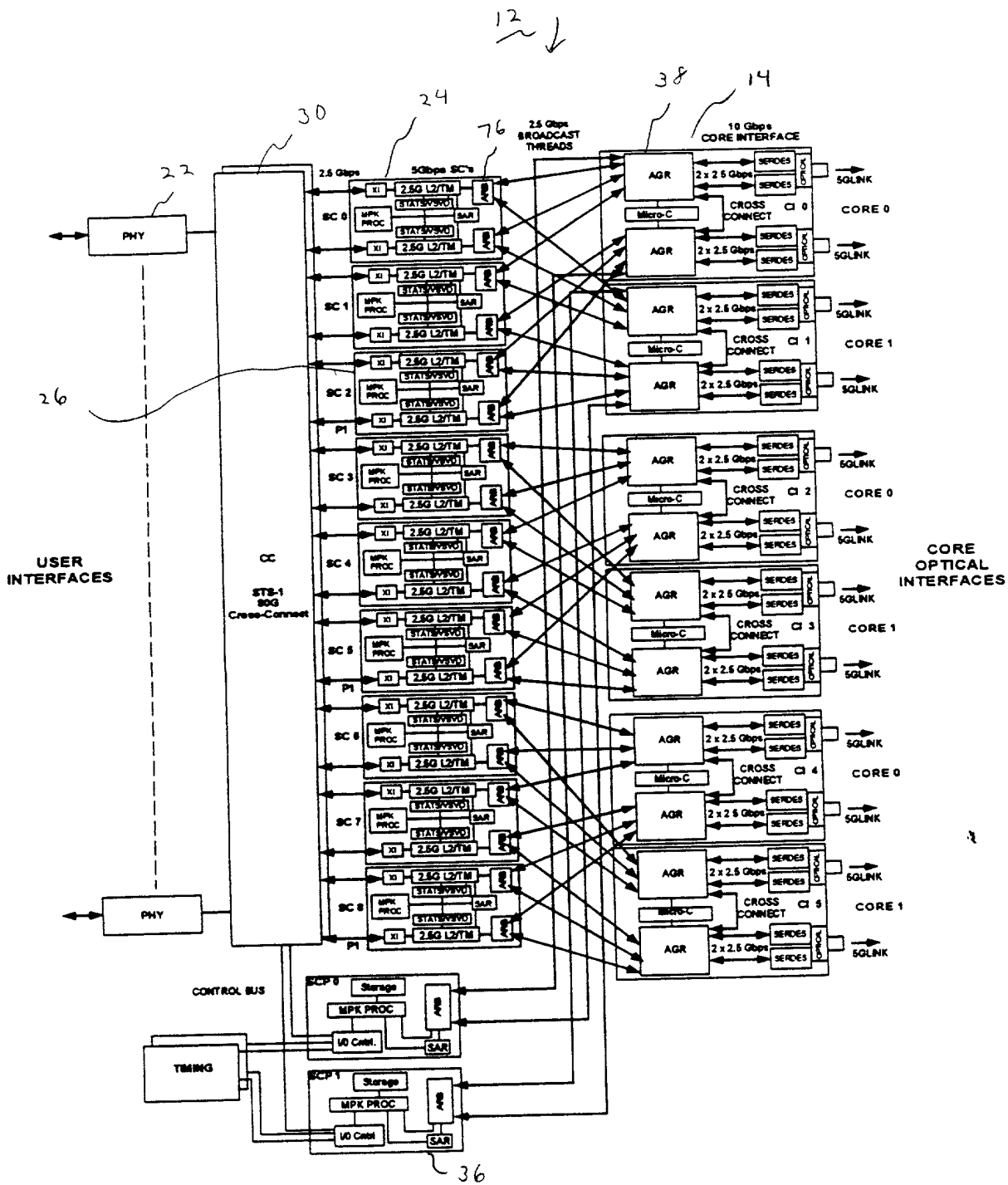


Fig 2



3/5

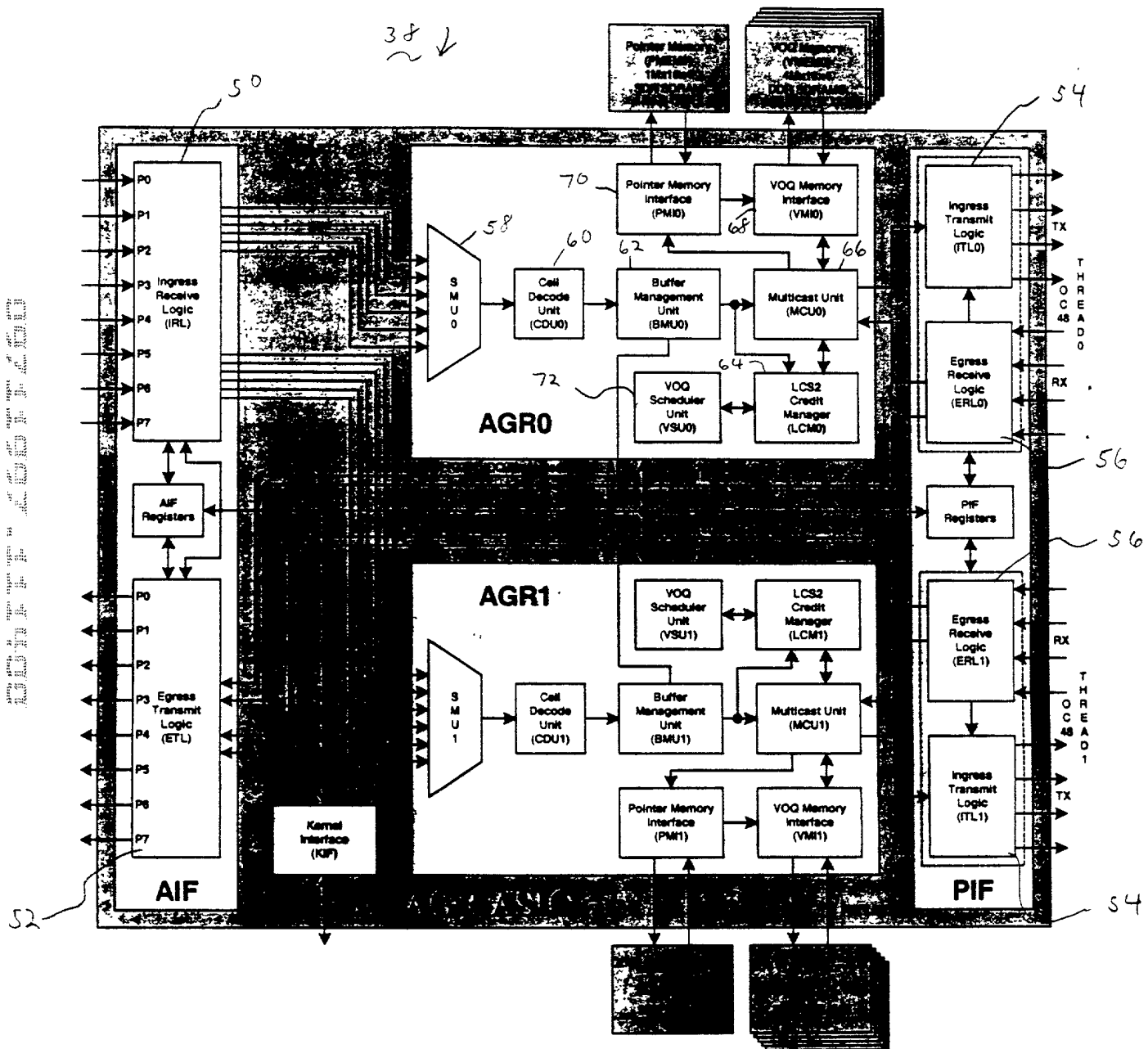
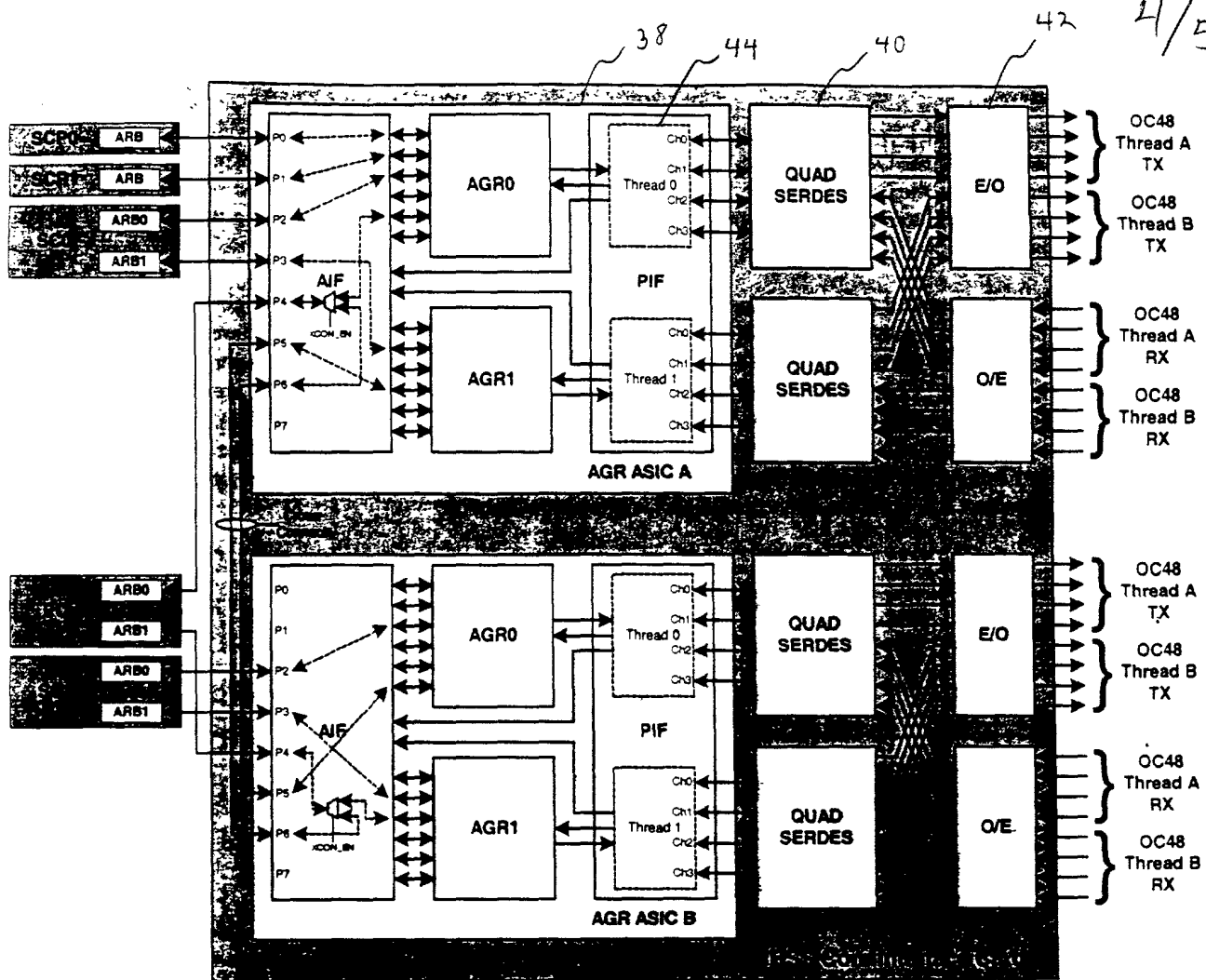


Fig 4



14

Fig 3

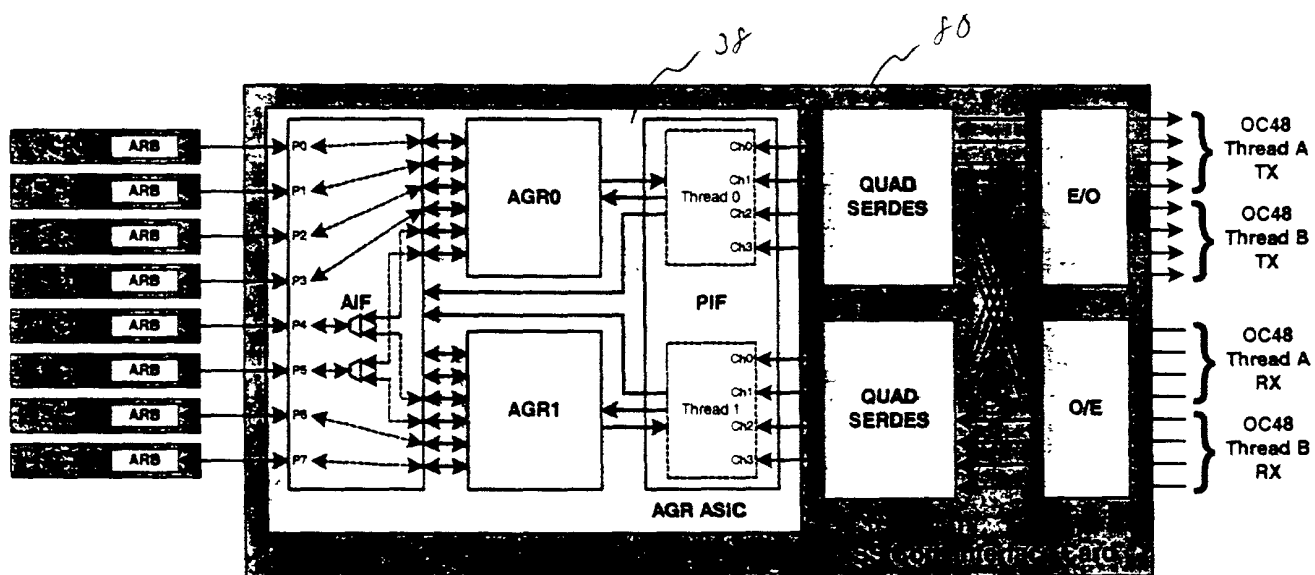


Fig 5

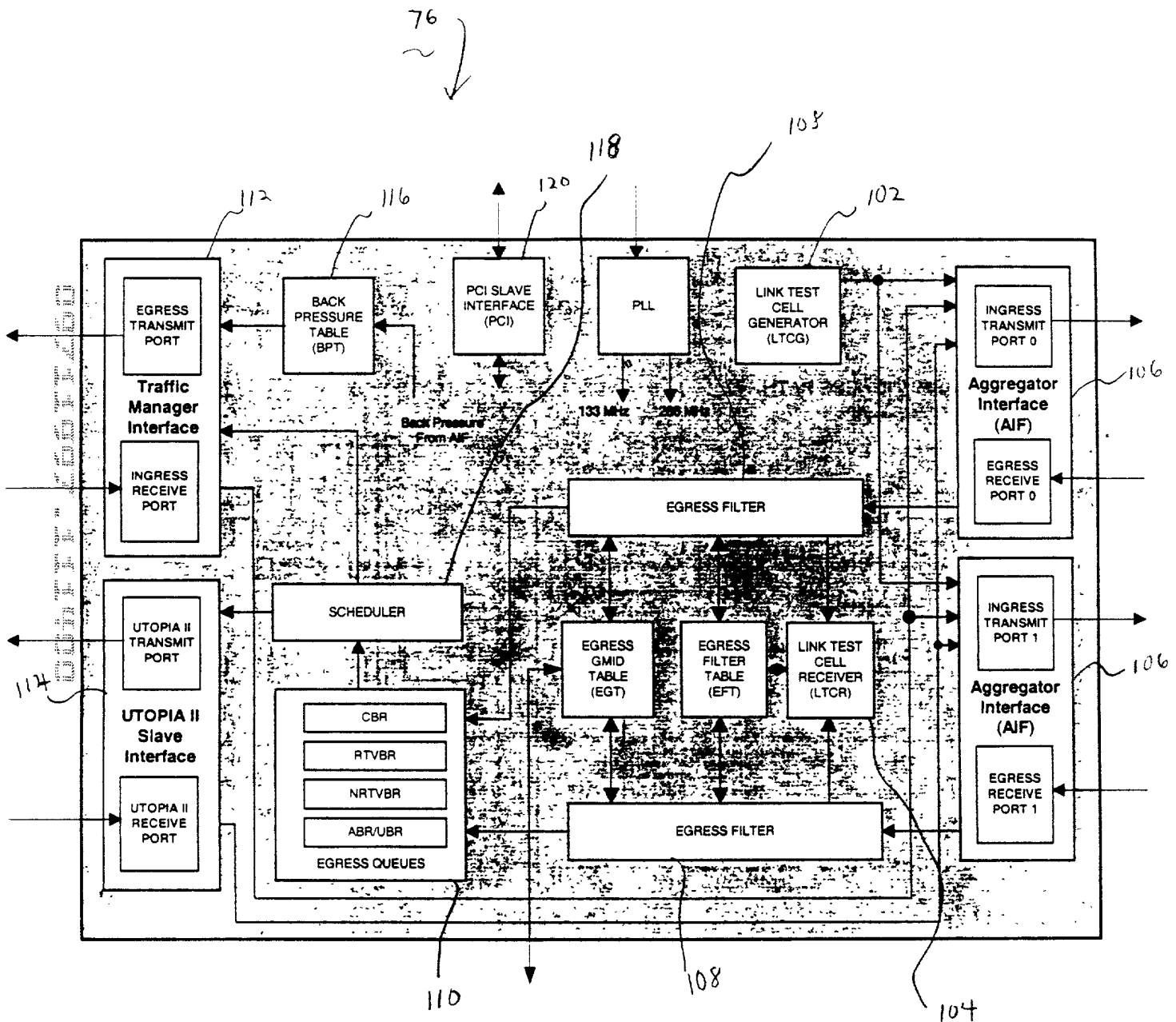


Fig 7

**IN THE UNITED STATES  
PATENT AND TRADEMARK OFFICE**

**Declaration and Power of Attorney**

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am an original, first and joint inventor of the subject matter which is claimed and for which a patent is sought on the invention entitled **APPARATUS AND METHOD FOR REDUNDANCY OF PROCESSING MODULES INTERFACED TO A SWITCHING CORE** the specification of which is attached hereto.

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by an amendment, if any, specifically referred to in this oath or declaration.

I acknowledge the duty to disclose all information known to me which is material to patentability as defined in Title 37, Code of Federal Regulations, 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, 119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

None

I hereby claim the benefit under Title 35, United States Code, 120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, 112, I acknowledge the duty to disclose all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, 1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application:

None

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

I hereby appoint the following attorney(s) with full power of substitution and revocation, to prosecute said application, to make alterations and amendments therein, to receive the patent, and to transact all business in the Patent and Trademark Office connected therewith:

Lester H. Birnbaum	(Reg. No. 25830)
Richard J. Botos	(Reg. No. 32016)
Jeffery J. Brosemer	(Reg. No. 36096)
Kenneth M. Brown	(Reg. No. 37590)
Donald P. Dinella	(Reg. No. 39961)
Martin I. Finston	(Reg. No. 31613)
William S. Francos	(Reg. No. 38456)
Barry H. Freedman	(Reg. No. 26166)
Julio A. Garceran	(Reg. No. 37138)
Jimmy Goo	(Reg. No. 36528)
Anthony Grillo	(Reg. No. 36535)
Stephen M. Gurey	(Reg. No. 27336)
John M. Harman	(Reg. No. 38173)
Matthew J. Hodulik	(Reg. No. 36164)
Michael B. Johannesen	(Reg. No. 35557)
Mark A. Kurisko	(Reg. No. 38944)
Irena Lager	(Reg. No. 39260)
John B. MacIntyre	(Reg. No. 41170)
Christopher N. Malvone	(Reg. No. 34866)
John F. McCabe	(Reg. No. 42854)
Scott W. McLellan	(Reg. No. 30776)
Michael A. Morra	(Reg. No. 28975)
Gregory J. Murgia	(Reg. No. 41209)
Claude R. Narcisse	(Reg. No. 38979)
Joseph J. Opalach	(Reg. No. 36229)
Neil R. Ormos	(Reg. No. 35309)
Jack R. Penrod	(Reg. No. 31864)
Gregory C. Ranieri	(Reg. No. 29695)
Scott J. Rittman	(Reg. No. 39010)
Ferdinand M. Romano	(Reg. No. 32752)
Eugene J. Rosenthal	(Reg. No. 36658)
Ronald D. Slusky	(Reg. No. 26585)
David L. Smith	(Reg. No. 30592)
Ozer M. N. Teitelbaum	(Reg. No. 36698)
John P. Veschi	(Reg. No. 39058)
Charles L. Warren	(Reg. No. 27407)
Eli Weiss	(Reg. No. 17765)

Please address all correspondence to the Docket Administrator (Rm. 3C-512), Lucent Technologies Inc., 600 Mountain Avenue, P. O. Box 636, Murray Hill, New Jersey 07974-0636. Telephone calls should be made to Matthew J Hodulik by dialing 732-949-9742.

Full name of 1st joint inventor: P K Chidambaran

Inventor's signature Chidambaran Date 11/09/2000

Residence: Santa Clara, Santa Clara County, California

Citizenship: India

Post Office Address: 3775 Flora Vista Avenue,  
#602  
Santa Clara, California, 95051

11/09/2000 11:09 AM P.K. Chidambaran

Inventor's signature \_\_\_\_\_

Date \_\_\_\_\_

Citizenship: United States of America

Post Office Address: 674 Old Harvard Road  
Boxborough, Massachusetts, 01719

Inventor's signature *John L. Jones* Date 11/4/00

Citizenship: United States of America

Post Office Address: 49 Vose Hill Road  
Westford, Massachusetts, 01886



Full name of 5<sup>th</sup> inventor: Andrew A Long

Inventor's signature Chen A. Li Date 11-1-2000

Residence: Northborough, Worcester County, Massachusetts

Citizenship: United States of America

Post Office Address: 14 Wheelwright Drive  
Northborough, Massachusetts, 01532

Full name of 6<sup>th</sup> inventor: Prasasth R Palnati

Inventor's signature Prasasth Palnati Date November 13, 2000

Residence: Westford, Middlesex County, Massachusetts

Citizenship: India

Post Office Address: 27 Chicory Road  
Westford, Massachusetts, 01886

U.S. Patent and Trademark Office  
Washington, D.C. 20503  
www.uspto.gov

Inventor's signature Ronald T. Wankun Date Nov 8, 2000

Citizenship: United States of America

Post Office Address: 328 Joseph Road  
Boxboro, Massachusetts, 01719

Full name of 8<sup>th</sup> inventor: Raymond J Schmidt

Inventor's signature

Date \_\_\_\_\_

Residence: Stoughton, Norfolk County, Massachusetts

Citizenship: United States of America

Post Office Address: 8 Anderson Road  
Stoughton, Massachusetts, 02072